

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

of

Peter-Pike J. Sloan

and

Michael F. Cohen

for

Interactive Horizon Mapping

ATTORNEY'S DOCKET NO. MS1-791US

INTERACTIVE HORIZON MAPPING

[0001] TECHNICAL FIELD

[0002] The present invention relates to computer graphics renderings, and more particularly to systems, methods, and program products for real time shadowing of bump mapped surfaces using horizon mapping in computer graphics renderings.

[0003] BACKGROUND

Computer graphics involves the rendering of images for display on a visual terminal of a computer. Rendering involves the translation of the three-dimensional (3D) geometry of graphical objects into picture elements (pixels) that are stored in a frame buffer. The contents of the frame buffer can then be displayed on the visual terminal or display device.

[0004] In a computer gaming, 3D animation is used to display a sequence of images interactively to the user of the computer game, which gives the illusion of motion in three-dimensional space. This interactivity allows a user to change a scene or the point of view in the scene. These on-demand changes initiated by the user require a rendering system that can create new images in real-time.

[0005] Computer games are one example of a 3D computer graphics system. Others include computer-aided design (CAD) and virtual reality ("VR"). A 3D graphics accelerator is typically used in these types of systems. The graphics hardware is used to receive off-loaded processing functions from a host processor to speed up system performance.

[0006] An application program executing on a host processor generates three-dimensional geometry data including information corresponding to points on the surface of a three-dimensional graphical object. These points are usable as vertices of polygons which, when connected, may be rendered to form a representation of the graphical object. The application program causes the host processor to transfer this 3-D geometry data to a graphics

accelerator, along with corresponding control and connectivity information. The graphics accelerator receives this stream of compressed 3-D geometry data and, for reasons of speed and diversity, renders a series a collection of flat encoded polygons, for example triangles, in a frame buffer for subsequent display on an attached display device. A polygonal system such as this has the advantage of being able to translate and project such objects quickly and efficiently. The process of connecting three-dimensional vertices into component polygons, the sum of which represents a graphical object, is referred to as "tiling".

[0007] Each graphical object to be rendered has a coordinate system that is local to the graphical object within the 3D computer graphics system. This coordinate system needs to be transformed from the coordinate system of the graphical object to the pixel display coordinates in the frame buffer. Once the geometry of the object has been transformed into screen coordinates, it is rasterized or broken into a set of pixel color values that are stored into the frame buffer.

[0008] Each corner or vertex of each polygon has associated therewith certain values for pixel parameters based on the graphical object itself and the environmental properties of the graphical object. Interpolation is used for each pixel in the polygon area to generate the pixel parameters. One such pixel parameter is texture. Texture on the polygon is suggested by means of a texture map to accomplish finer detail in the polygon.

[0009] In conventional commodity graphics hardware, the surface of an object is typically digitally represented by a texture in terms of four (4) color channels, namely red, green, blue, and alpha. At each vertex of each polygon, two-dimensional (2D) coordinates are specified for the orientation in the tangent plane of the light. The 2D qualities include color and brightness and are encoded by the first three channels. A texture is also encoded with three-dimensional properties, such as how transparent and reflective the object is by use of

the fourth channel - the alpha channel. Once a texture is parameterized by a point on a surface, another texture can be parameterized by where the light is relative to the local surface point. Then, once a texture has been defined, it can be wrapped around any 3-dimensional object in a texture mapping process.

[0010] A texture map is a picture stored in a computer memory as a bit map that is applied to the polygons that make up the graphical object to make the graphical object look more realistic. The texture map is comprised of individual two dimensional space rectangular 'texels' coordinates that depict a certain texture. When the texels are mapped onto a polygon, it will appear to be textured.

[0011] Texture mapping processes can be used to render an appearance of a smooth surface. A more realistic and topographical rendering, however, can be accomplished using a texture map that is known as a bump map. Bump mapping uses a bump map that is associated with a surface. The bump map, which is normally pre-calculated into a look up table stored in memory, is a collection of different surface normal vectors on different surface locations of the surface to create an impression of a rough surface. Stated otherwise, bump mapping simulates the effect of light reflecting from small perturbations across a surface. Bump mapping, often considered as the application of a shading function to an object, is done on a pixel-by-pixel basis.

[0012] Bump mapping techniques, which convey surface texture by perturbing the normal vectors of a surface, are available on conventional commodity graphics hardware. The advantage of bump mapping is that it provides a simple way to modulate the shading on a surface. Bump maps represent a special kind of texture map. Instead of a color, each texel of a bump map contains a height field gradient. Bump maps are applied to a surface of an object using standard texture-mapping tools. Very complex surfaces (wrinkles, folds,

bumps) can be simulated without increasing the complexity of the surface itself. A space in which the bump map is defined is called a map space.

[0013] In bump mapping, texel coordinates, designated as “u” and “v”, are used in a single component texture map, $f(u, v)$. The texture map is interpreted as a height field that perturbs at each point on the surface along its normal vector $N = (P_u \times P_v) / |P_u \times P_v|$, (1)

where P_u and P_v are partials of the parameterization of texture space axes u and v. Rather than actually changing the surface geometry, however, only the normal vector is modified.

[0014] Shadows play an important providing perceptual cues for perceiving surface shape and surface texture of an object. Bump mapping, however, does not define any explicit geometry such that there are no actual bumps to cast shadows. Stated otherwise, a bump defined by a conventional bump mapping function can cast neither a bumpy shadow nor a shadow on another object. Thus, interactive shadowing techniques that rely on an explicit representation of the geometry cannot be used to simulate shadows cast by the virtual bumps implied in bump mapping.

[0015] It would be an advance in the art to provide a technique to render with commodity graphics hardware, in real time at interactive rates, a graphical image of an object having a bump mapped surface, where the bumps on the bump mapped surface cast shadows, thus more accurately accounting for the geometry of the underlying bump mapped surface. It would also be an advance in the art to represent limited self-shadowing for pure geometry while maintaining interactivity with a user of a computer.

[0016] SUMMARY

[0017] Interactive, real time self-shadowing, also known as horizon mapping, of a bump mapped surface is implemented for computer rendered objects. To do so, a system, method and program product are provided to represent a computer graphics textured object function that is defined over an orientation in a tangent space of the object using different textures or basis functions. The implementations described herein can be performed using commodity graphics hardware in real time, interactively.

[0018] A technique is presented for casting of shadows by a precomputation of limited visibility for each point on a bump mapped surface. In particular, the lowest angle to the horizon is encoded in a discrete number of radial directions to each point to represent at what height each point becomes visible from each radial direction. This parameterization produces the self-shadowing of the surface geometry.

[0019] BRIEF DESCRIPTION OF THE DRAWINGS

[0020] Generally, the same numbers are used throughout the drawings to reference like elements and features.

[0021] Figure 1 depicts a bumped curved surface having points thereon and an angle therefrom to the lowest eastern horizon.

[0022] Figure 2 depicts one of the points seen in Figure 1 at the origin, a projection of a vector to the light source upon the tangent plane contain the point in Figure 1, and an angle that represents the orientation of the projected vector in the tangent plane.

[0023] Figure 3 depicts a normal N taken from a plane that is perpendicular to the surface seen in Figure 1 at the point seen in Figures 1 and 2.

[0024] Figure 4 is a flow diagram showing an illustrative methodological implementation of the invention herein, and particularly showing preprocess precomputation operations performed noninteractively in texture space.

[0025] Figure 5 is a flow diagram showing run time operations to follow the precomputation operations of Figure 4.

[0026] Figure 6 is a flow diagram showing run time operations to follow the run time operations of Figure 5.

[0027] Figure 7 is a flow diagram showing run time operations to follow the run time operations of Figure 6.

[0028] Figure 8 is an example of a computing operating environment capable of implementing, either wholly or partially, an illustrative implementation of the invention.

[0029] Figures 9A-9B through Figure 17A –17C are renderings, in various stages thereof, that were produced using the interactive horizon mapping implementations disclosed herein.

DETAILED DESCRIPTION

[0031] Bump and Horizon Mapping

In dot product bump mapping, also known normal mapping or bump mapping, vectors that are normal (“the normals”) to a surface are modified instead of modifying the height of the surface. The modifications to the normals render diffuse and specular reflections from small surface bumps on the surface. The technique directly stores the normals of the surface in texture maps. Bump mapping is particularly effective in rendering visual cues because the prime component in the calculation of the intensity of a picture element (pixel) is the direction of the surface normal at that pixel. For each point on a surface, there corresponds a height value by which this point would be raised on the surface.

[0032] Horizon mapping approximates the shadows cast by the bumps on the same surface. As such, a computation can be made to self-shadow bump maps by describing the horizon for a small number of directions at each point in the height field. During rendering a shadow test is performed to determine, for each point on the height field, whether the light direction is above or below the interpolated horizon.

[0033] The bump function produces a two dimensional table of height values for each point on the height field. A computation is made and stored, for each of the height values, of the lowest angle between a vector towards the visible horizon and the surface plane. A series of directions around the horizon, such as at the eight (8) main compass points - e.g. N, S, E, W, NE, SE, SW, and NW), are computed for each point on the height field of the surface plane.

[0034] During the rendering stage, the lowest horizon angle at the intersection point is interpolated from the light direction and the horizon map for each point on the height field. If the lowest horizon angle from the surface normal is greater than the angle from the light direction to a point on the height field point, then that point is in shadow. Shadowing of bump mapped surfaces is implemented in an algorithm that maximizes the power of a frame

of the surface P_u and P_v . Given a bump map $F(u,v)$ as a non-negative scalar function parameterized over the same domain, the surface normal N can be modified, after dropping terms of first order, as follows:

$$P'_u = P_u + F_u N / |N|, P'_v = P_v + F_v N / |N|, N' = P'_u \times P'_v = N + D, \quad (2)$$

$$\text{where } D = (F_u N \times P_v - F_v N \times P_u) / |N| \quad (3)$$

is the perturbation of the normal.

[0039] Horizon mapping is performed into the local coordinate system of the surface. This mapping is performed through the dual of the basis P_u, P_v, N , where N is the unbumpmapped surface normal which can be determined by an affine transform. The mapping is computed by building a matrix with the basis as columns and inverting it, where the rows of this inverse are a scaled version of $N \times P_v$ and $P_u \times N$, along with N itself. A vector in this local frame has an associated orientation in the tangent plane θ and an angle with the normal ϕ .

[0040] There is then a determination made, at each pixel, whether a light vector, when transformed into the local coordinate frame, is visible above the horizon. To discover this, a horizon map, $\phi_{u,v,\theta}$, is built. The horizon map is tabulated at discrete u, v parameter values, and in a set of directions θ , to represent the azimuth angle ϕ when a light would become visible at each direction θ . By way of example, u and v are sampled fairly densely (e.g. 512 x 512), and θ is sampled more coarsely (the eight (8) main compass points - e.g. N, S, E, W, NE, SE, SW, and NW).

[0041] An advantage is gained in horizon mapping of bumped, curved surfaces. Particularly, the precomputed horizon map is created in terms of the local geometry at each discrete (u, v) coordinate. This is possible and practical in the present implementation because the underlying geometry of the bumped, curved surface is known at the time when the horizon map is computed.

[0042] Given the variable M to represent a number of sampled directions (the eight (8) main compass points - e.g. N, S, E, W, NE, SE, SW, and NW) for θ and the discrete domain coordinates u_i, v_j , then the horizon angle $\phi(u_i, v_j, \theta)$ for any direction at coordinates u_i, v_j is found by interpolating between the discrete directions as follows:

$$\phi(u_i, v_j, \theta) = \sum_{k=1}^M B_k(\theta) \phi(i, j, k), \quad (4)$$

where $B_k(\theta)$ is a basis function for the k^{th} direction. The basis function $B_k(\theta)$ in the linear case is evaluated to one for the corresponding direction and linearly falls off to zero for the neighboring directions. For example, if the k th direction is east then the value of the linear basis function $B_k(\theta)$ is one when θ equals east, and the value falls off to zero at the two closest compass points (northeast (NE) and southeast (SE)). In other words, the horizon angle ϕ is interpolated between discrete directions, noting that since this is a radial function it wraps around to interpolate between $\phi(i, j, 1)$ and $\phi(i, j, M)$. Similarly, the function $\phi(u, v, \theta)$ is bilinearly interpolated across the parameters u and v . In general the basis functions can be any functions that are all positive and sum to 1 for any orientation θ . The coefficients for linear basis functions can be determined just by point sampling a discrete number of directions, but in general the coefficients for any shape of basis functions can be computed by solving a simple linear least squares problem that relates the coefficients of the basis functions to a potentially super sampled set of horizon angles at various orientations.

[0043] To test the visibility of a light source from a point on a surface, the direction (x, y, z) from the point to the light source is transformed into the local tangent frame by a simple affine transform to provide the local coordinates (u, v, θ) . While the angle θ could be determined by projecting the transformed vector onto the tangent plane, it is desirable that the first two transformed coordinates are used as a lookup into a table of θ directions. The

look up table will, by example, be encoded as a texture in the computer hardware. The table may be considered to be a set of texture maps.

[0044] The texture mapping feature is conventional in commodity graphics hardware. By using the texture mapping feature that is provided with conventional commodity graphics hardware to store horizon maps, the need for sophisticated hardware and expensive host processor time is avoided. As such, horizon mapping of curved, bumped surfaces is enabled using the texture mapping capabilities of conventional commodity graphics hardware.

[0045] General Self-Shadowing with Horizon Maps

[0046] The tangent plane parameterization and discretization used to compute shadows for bump maps can also be used to store global visibility information about a surface in general. The horizon angles are determined by shooting rays into a graphical object, starting from the horizon and moving towards the surface normal until there is no intersection. The largest angle that can be represented, by this determination of the horizon angles in the horizon map, is ninety (90) degrees out of the tangent plane. This determination of the shadows from bump maps is sufficient for most surfaces that lack undercut features.

[0047] Implementation On Commodity Graphics Hardware

[0048] As was generally discussed above, horizon mapping can be performed for curved, bumped surfaces using the texture mapping capabilities of conventional commodity graphics hardware. A discussion is now presented, in reference to Figures 4-7, for both the precomputation and the runtime shadowing algorithm and implementation thereof for commodity graphics hardware.

[0049] Precomputation

[0050] The surface to be bumped and horizon mapped is tiled into a collection of polygons to which a texture is to be mapped in a parameterization thereof. The bump map will be

represented in one such texture parameterization of the surface. The parameterization, seen at Block 402 in Figure 4, dictates how the bump map will wrap or fit upon the surface. The algorithm requires several inputs, including the coordinate values u, v associated with every vertex of each polygon, the geometry of the surface as a parameterization in the u, v coordinates, and a non-negative scalar valued bump map, $F(u, v)$, depicted at Block 404, that is parameterized over the surface.

[0051] Before a graphical object is displayed upon a peripheral device, such as a computer screen, so as to be visible to a computer user, precomputation operations first are performed upon the inputs discussed above. For each vertex of each polygon representing a surface on each graphical object, a vector valued perturbed normal map, $N'(u, v)$ is precomputed at Block 406 from the non-negative scalar valued bump map, $F(u, v)$. The normal map carries the bump map into the normal at every pixel. Stated otherwise, the surface height is perturbed at every pixel in this precomputation phase.

[0052] To compute the normal map in texture space, there are given 'M' directions in the tangent plane for an angle $\theta_{k=1..M}$. M will, by example, be represented by the eight (8) main compass points (e.g. N, S, E, W, NE, SE, SW, and NW). As such, one horizon map, $\phi(u, v, \theta_k)$, is computed for each θ_k at Block 408, for a total of M horizon maps. When M=8, the eight horizon maps be collected into only two (2) horizon maps by encoding four (4) of the eight directions into the four (4) color channels of texture storage. It is intended that the first direction be stored in the red channel, the second direction be stored in the green channel, the third direction be stored in the blue channel, and the forth direction is stored in the alpha (α) channel. As such, the first horizon map might encode the directions N, NE, E and SE, while a second map encodes the directions S, SW, W and NW. Stated otherwise, a plurality of M radial directions could be approximately equally distributed from about zero

radians to about 2π radians about each coordinate point (u, v) . By way of example, one radial direction could be taken from about zero radians to about $\pi/2$ radians, another radial direction could be from about $\pi/2$ radians to about π radians, another radial direction could be from about π radians to about $3/2\pi$ radians, and still another radial direction could be from about $3/2\pi$ radians to about 2π radians.

[0053] M basis maps are also precomputed. As such, there is one (1) basis function for each of the eight (8) discrete directions. Each of the M basis maps are derived at Block 410 and are represented by $B_k(s, t)$, where (s, t) represents the influence of the direction of the light source, θ_k , as set forth in Eq. 4, above. The encoding of the M basis maps will, by example, be similar to the encoding of the M horizon maps, again using the texture capabilities of conventional commodity graphics hardware. Accordingly, two (2) basis maps are encoded to each contain four (4) of the eight (8) directions.

[0054] The encoding scheme for the normal and horizon maps realize a computational efficiency in that the dot products can be computed four (4) at a time by multiplying the weight of the basis function (the orientation value) times the horizon values. This implementation represents functions over orientations on a surface such that, for every point, for any orientation, there will be some function.

[0055] An optional precomputation, seen at Block 412, can derive a one dimensional arccosine mapping for each mapped point from the cosine of the horizon angle ($\cos(\phi)$) to the horizon angle ϕ . This precomputation compares the cosines of the horizon angles instead of comparing the horizon angles themselves. Such an approximation is more readily realized because the dot products yield cosines.

[0056] Another precomputation operation that is performed for each vertex is the scaling of the duals or a dual for the tangent frame basis, where computations are made for

transformations into texture space. This per vertex computation, which is seen at Block 414, is represented as

$$[P_u, P_v, N]^{-1} = [S^T, T^T, N^T], \quad (5)$$

and inverts the non-bump mapped local tangent frame. For a planar surface, the local tangent plane is the same for all vertices, but varies at each vertex over a curved surface. The inversion of the non-bump mapped local tangent frame allows a ready transform of the light direction onto the local tangent plane of the surface at each frame time.

[0057] These precomputed data could be stored in volatile or nonvolatile memory since the data will not be needed until run time. By way of example and not by way of limitation, computer gaming applications may store horizon maps in precomputed texture space on computer readable media such as a CD-ROM. When the computer game is played, the CD-ROM is read by a computer having conventional commodity graphics hardware, or its functional equivalent, to access and use the precomputed horizon maps.

[0058] At Each Frame Time, At Run Time

[0059] Use is made of the precomputed data during the interactive execution of the associated graphics program. For each scene to be rendered, given a current lighting condition, a computation is made at each vertex of each polygon for each graphical object for (s, t, ϕ_L) , where s, t are the orientation coordinates of a point in the tangent plane at the vertex, and ϕ_L is the angle between the light and the normal at that point. The view to be rendered at this point is not taken into account because the shadows to be depicted are not view-dependent.

[0060] Set Up Camera To Draw Into Texture Space

[0061] A computation is made at Block 502 in Figure 5 for each point as to the horizon angle that the light source would have be greater than for that point not to be in the shadow

generated by the topography from that light source. Given the precomputation above, at each frame time, light is first projected in a light direction onto the local tangent plane at each vertex of each polygon. The light vector when dotted with the first two components of the inverted frame S and T yields the projection of the light vector into the coordinate space in the tangent plane resulting in the pair, (s, t) . The light vector dotted with the normal at each vertex gives $\cos(\phi_L)$.

[0062] The remainder of the computation is carried out per pixel and is done in hardware using multi-texturing and blending into the frame buffer. The multi-texturing that is used in this implementation is different than conventional multi-texturing which multiplies the normal map times the diffuse color. This implementation, however, sums up the weighted basis functions instead.

[0063] The multi-texturing begins at Block 504 by setting transformations so as to render into UV space, by using the u, v coordinates as vertex coordinates. To do so, the textures are multiplied by the components thereof, the results of the multiplications are summed, and the results are stored in the alpha (α) channel. The contribution for each of the first four (4) of eight (8) directions are accumulated into the frame buffer at Block 506 in Figure 5. The first texture in the multi-texturing process is the basis map $B_1(s, t)$ for the directions east, northeast, north, and northwest, where the s, t coordinates are for each vertex and are derived from the light direction. The second texture is the horizon map, $\phi_1(u, v)$, for the same directions as the first texture, but indexed by the u, v from the surface.

[0064] The contribution for each of the last four (4) of eight (8) directions are accumulated into the frame buffer at Block 508. The first texture is the basis map $B_2(s, t)$ for the directions east, northeast, north, and northwest, and the second texture is the horizon map, $\phi_2(u, v)$, for the same directions as the first texture. The resulting alpha (α) channel now

represents the horizon angle, ϕ , in the direction of the light. The results are saved into a texture map represented by ϕ (θ_{LIGHT}).

[0065] The foregoing basis function and horizon map computations can be multiplied together and then added into the frame buffer for each of the eight (8) directions. The frame buffer then stores, for every pixel on each surface and for each of the eight directions, the angle to the horizon for each corresponding point on the surface.

[0066] Set Up Camera To Draw Into Scene

Following the storage of the horizon angles in the frame buffer at Block 510 in Figure 5, three (3) rendering passes are performed, each computing the transformations to draw into the current view of camera of the scene. The first pass, seen in Figure 5 at reference numeral 512, draws the model with only ambient lighting in the frame buffer. Also, the depth buffer is initialized to compute the distance of every pixel from the front of the scene such that only the front most pixel is visible.

[0067] A second of the three passes, seen in Figure 6 at reference numeral 602, creates a stencil that will only allow non-shadowed pixels to be rendered. This pass sets an alpha test, at Block 604, which will accept pixels for display only if the corresponding alpha channel value is non-zero. All pixels having a non-zero value in the corresponding alpha channel are permitted to pass through at Block 606. As such, the graphical object to be rendered is drawn to avoid drawing any point that is in shadow. Procedurally, no point is drawn where the stencil has a bit corresponding to the point that is set so as to be representative of the point being in shadow. Then, iterations are performed for each of the eight (8) directions.

[0068] After the directional iterations are computed in the second pass, a color mask is then set at Block 608 so that color channels are not written. The color mask thus preserves the ambient term that was set in the first pass.

[0069] Following the color mask operation, a multi-texturing function is performed. The multi-texturing function is set to subtract values at Block 610, where any negative value is set to zero, in that negative numbers are not meaningful to shadow renderings.

[0070] The scene is drawn at Block 612 in Figure 6 using the precomputed two (2) sets of texture coordinates. These texture coordinates represent, for each point given the direction of the current light source, the angles from the tangent plane of the point to the light source and the angles from the surface normal to the lowest horizontal point on the topography. Subtracting the two texture values will yield a light-value in the alpha (α) channel. Positive light-values will result in the surface being drawn in the light and zero values will result in the surface being drawn in shadow, in that negative values are set to zero in the alpha channel.

[0071] The two textures used in the second pass and are represented as a 1D texture and a 2D texture. The 1D texture is $\cos(\phi) > \phi$, where $\cos(\phi)$ of the light was previously computed and store for at each vertex. This result contains the angle off the normal to the light at each pixel. The 2D texture is ϕ (θ_{LIGHT}) that was previously computed and stored for each vertex. The 2D texture encodes the angle off the normal of the horizon in the direction of the light at each pixel.

[0072] In essence, the foregoing 1D and 2D texture operations make use of the red, blue, green, and alpha channels that are conventionally used for lighting in the texture feature of commodity graphics hardware. The present implementation, however, uses these color channels to combine and weight the horizon textures with the basis textures in a reconstruction of the texturing function. As such, the implementation uses the general notion of a dot product different than is conventional. Accordingly, less computations are performed for each of the eight (8) directions, in that only two (2) passes are required,

namely four (4) channels and directions on the first pass and four (4) channels and directions on the second pass, resulting in a computation that is four times as fast.

[0073] The last of the three passes, the initiation of which is seen in Figure 7 at reference numeral 702, performs a normal bump mapped rendering of the model wherever the pixels in the frame buffer are not in shadow. Since a test of the alpha channel would not be meaningful where points are in shadow, the alpha test is turned off at Block 704. A stencil is set at Block 706 to prevent any pixel not having its stencil bit set from being drawn so that only those pixels intended to be seen are drawn. Then, at Block 708, a color mask is set in place to allow for the writing of the color channels into the frame buffer. A blending function is enabled at Block 710. The bending function adds other lighting to the ambient low level lighting that was previously accounted for in the frame buffer. The other lighting to be added to the image in this blending function includes diffuse and specular lighting.

[0074] Following the blending of other lighting, previously computed normal maps are added into the frame buffer at Block 712 to the graphical object to be rendered from the contents of the frame buffer in a standard bump mapping operation. Block 712 has the effect of shading the non-shadowed regions of the graphical object given the lighting present on the surfaces being rendered. Optionally, methods other than conventional normal mapping over the surfaces could also be used, such as bit mapped textures, etc.

[0075] Following the last of the three (3) passes, an image is rendered on a display device at Block 714 so as to display an ambient only term in shadowed regions and a normal bump mapping in non-shadowed regions. In order to create lighter shadows, the shadowed regions can be rendered with a lighter diffuse term rather than ambient only. This can be done in a fourth pass, not shown, by first setting the stencil function to only draw pixels that do not have the stencil bit set and by then drawing the geometry again using normal bump

mapping, but with the scaled-down diffuse term. This fourth pass can also be combined with the first ambient pass. Iterating on creating the light dependent horizon map and the resulting alpha test can also be done for multiple light sources.

[0076] The results of the foregoing implementation, in various stages of renderings, are depicted in Figures 9A-9B through Figures 17A-17C.

[0077] Figure 9A depicts a basis function for the bump mapped surface rendering that is self-shadowed by horizon mapping and seen in Figure 9B.

[0078] Figure 10A shows the rendering of a normal map of a surface which is seen in Figure 10B with a horizon mapping that includes a light source coming from the north, and where Figure 10C illustrates the northerly light source basis texture for Figure 10B.

[0079] Figure 11A is a rendering of a surface that is bump mapped, where the surface is rendered in Figure 11B with shadowing via horizon mapping by a northeasterly light source.

[0080] Figures 12A and 12B, both of which are seen as being lit by a northeasterly light source, are renderings that feature horizon mapping techniques that are, respectively, point sampling and curve fitting.

[0081] Figures 13A and 13B are renderings of the same bump mapped surface lit by a southwesterly light source, where Figure 13A illustrates shadows rendered by point sampled horizon mapping, where Figure 13B features the shadow that are smoother than that of Figure 13A through the use of a super sampling of curve fitting in a horizon mapping technique.

[0082] Figures 14A and 14B are renderings of the same bump mapped surface lit by a southwesterly light source, where Figure 14A has rougher shadows of point sampled horizon mapping and Figure 14B features smoother shadows that are accomplished by sampling more points than the rendering seen in Figure 14A.

[0083] Figures 15A-C are renderings of bump mapped surface of a simple plane, and respectively showing a normal map without shadowing, dense shadows accomplished by horizon mapping, and light shadows accomplished by horizon mapping.

[0084] Figures 16A-C are renderings of the same bump mapped cylindrical surface having 160 vertices and 158 triangles, and respectively showing the cylindrical surface without horizon mapping, with dense shadows accomplished by horizon mapping, and with light shadows accomplished by horizon mapping.

[0085] Figures 17A-C are renderings of the same bump mapped tessellated BSpline surface having 1600 vertices and 3042 triangles, and respectively showing dense shadows accomplished by horizon mapping, light shadows accomplished by horizon mapping, and light from a different direction seen from a different view, where Figures 17A and 17B simulate a shadow from a raised surface that is cast upon a shadow from the recessed surface.

[0086] It may be advantageous to alter the conventional graphics pipeline processing for the benefit of the foregoing implementation so as to reduce the number of passes required. One such option would be to put the alpha and stencil tests after the blending of fragments with the color already in the frame buffer. This alteration would eliminate the frame-buffer-to-texture-memory copy operation. To do so, the lighting direction could be written into tangent space into the alpha channel when the ambient term is written. Then, subtractions are made from the contributions from each of the direction passes. This would also eliminate one rendering pass for the geometry.

[0087] Having a versatile vertex shader, other than the functions of same that have been described herein, would allow the precomputation of the s , t texture coordinates, rather than requiring computation of the same at run time. It would be desirable that such versatility be

fully interoperable with the transform and lighting components of conventional commodity graphics hardware.

[0088] As an option to the above described implementation a different implementation in commodity graphics hardware can also be made without leveraging dot product fragment operations. It has been found, however, that a significant performance penalty will be had without use of dot product fragments.

[0089] The inventors intend these exemplary implementations to be examples and not to limit the scope of the present invention. Rather, the inventors have contemplated that the present invention might also be embodied and implemented in other ways, in conjunction with other present or future technologies

[0090] Exemplary Computing System and Environment

Figure 8 illustrates an example of a suitable computing environment 800 within which interactive horizon mapping, as described herein, may be implemented (either fully or partially). The computing environment 800 may be utilized in the computer and network architectures described herein.

[0091] The exemplary computing environment 800 is only one example of a computing environment and is not intended to suggest any limitation as to the scope of use or functionality of the computer and network architectures. Neither should the computing environment 800 be interpreted as having any dependency or requirement relating to any one or combination of components illustrated in the exemplary computing environment 800.

[0092] Interactive horizon mapping may be implemented with numerous other general purpose or special purpose computing system environments or configurations. Examples of well known computing systems, environments, and/or configurations that may be suitable for use include, but are not limited to, personal computers, server computers, thin clients,

thick clients, hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

[0093] Interactive horizon mapping may be described in the general context of computer-executable instructions, such as program modules, being executed by a computer.

Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types.

Interactive horizon mapping may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located in both local and remote computer storage media including memory storage devices.

[0094] The computing environment 800 includes a general-purpose computing device in the form of a computer 802. The components of computer 802 can include, by are not limited to, one or more processors or processing units 804, a system memory 906, and a system bus 908 that couples various system components including the processor 904 to the system memory 906.

[0095] The system bus 908 represents one or more of any of several types of bus structures, including a memory bus or memory controller, a peripheral bus, an accelerated graphics port, and a processor or local bus using any of a variety of bus architectures. By way of example, such architectures can include an Industry Standard Architecture (ISA) bus, a Micro Channel Architecture (MCA) bus, an Enhanced ISA (EISA) bus, a Video Electronics

Standards Association (VESA) local bus, and a Peripheral Component Interconnects (PCI) bus also known as a Mezzanine bus.

[0096] Computer 802 typically includes a variety of computer readable media. Such media can be any available media that is accessible by computer 802 and includes both volatile and non-volatile media, removable and non-removable media.

[0097] The system memory 806 includes computer readable media in the form of volatile memory, such as random access memory (RAM) 810, and/or non-volatile memory, such as read only memory (ROM) 812. A basic input/output system (BIOS) 814, containing the basic routines that help to transfer information between elements within computer 802, such as during start-up, is stored in ROM 812. RAM 810 typically contains data and/or program modules that are immediately accessible to and/or presently operated on by the processing unit 804. System memory 806 is an example of a means for storing data having inputs and outputs and a frame buffer for storing pixel representations from which to render a three-dimensional graphical object.

[0098] Computer 802 may also include other removable/non-removable, volatile/non-volatile computer storage media. By way of example, Figure 13 illustrates a hard disk drive 816 for reading from and writing to a non-removable, non-volatile magnetic media (not shown), a magnetic disk drive 818 for reading from and writing to a removable, non-volatile magnetic disk 820 (e.g., a “floppy disk”), and an optical disk drive 822 for reading from and/or writing to a removable, non-volatile optical disk 824 such as a CD-ROM, DVD-ROM, or other optical media. The hard disk drive 816, magnetic disk drive 818, and optical disk drive 822 are each connected to the system bus 808 by one or more data media interfaces 826. Alternatively, the hard disk drive 816, magnetic disk drive 818, and optical

disk drive 822 can be connected to the system bus 808 by one or more interfaces (not shown).

[0099] The disk drives and their associated computer-readable media provide non-volatile storage of computer readable instructions, data structures, program modules, and other data for computer 802. Although the example illustrates a hard disk 816, a removable magnetic disk 820, and a removable optical disk 824, it is to be appreciated that other types of computer readable media which can store data that is accessible by a computer, such as magnetic cassettes or other magnetic storage devices, flash memory cards, CD-ROM, digital versatile disks (DVD) or other optical storage, random access memories (RAM), read only memories (ROM), electrically erasable programmable read-only memory (EEPROM), and the like, can also be utilized to implement the exemplary computing system and environment.

[0100] Any number of program modules can be stored on the hard disk 816, magnetic disk 820, optical disk 824, ROM 812, and/or RAM 810, including by way of example, an operating system 826, one or more graphics application programs 828, other program modules 830, and program data 832. Each of such operating system 826, one or more graphics application programs 828, other program modules 830, and program data 832 (or some combination thereof) may include an embodiment of program code to perform interactive horizon mapping.

[0101] A user can enter commands and information into computer 802 via input devices such as a keyboard 834 and a pointing device 836 (e.g., a “mouse”). Other input devices 838 (not shown specifically) may include a microphone, joystick, game pad, satellite dish, serial port, scanner, and/or the like. These and other input devices are connected to the processing unit 804 via input/output interfaces 840 that are coupled to the system bus 808,

but may be connected by other interface and bus structures, such as a parallel port, game port, or a universal serial bus (USB).

[0102] A monitor 842 or other type of display device can also be connected to the system bus 808 via an interface, such as a video adapter/accelerator 844. Video adapter/accelerator 844 is intended to have a component thereof that represents 3-D commodity graphics hardware. As such, the 3-D commodity graphics hardware is coupled to the high-speed system bus 806. The 3-D commodity graphics hardware may be coupled to the system bus 808 by, for example, a cross bar switch or other bus connectivity logic. It is assumed that various other peripheral devices, or other buses, may be connected to the high-speed system bus 808, as is well known in the art. Further, the 3-D commodity graphics hardware may be coupled through one or more other buses to system bus 808.

[0103] Host processor 804 may transfer information to and from the 3-D commodity graphics hardware according to a programmed input/output (I/O) protocol over the system bus 808. In one implementation, the 3-D commodity graphics hardware accesses the system memory 806 according to a direct memory access (DMA) protocol, or through intelligent bus mastering.

[0104] In one implementation, one or more graphics application programs 828 executing on processing unit 804 generates three-dimensional geometry data. The resulting 3-D geometry data can be stored in system memory 806 and then transferred to the 3-D commodity graphics hardware over system bus 808. The 3-D commodity graphics hardware renders the resulting primitives accordingly on monitor 842. The 3-D geometry data may also be transferred over a network for subsequent rendering, or stored to a removable medium such as a CD-ROM. The 3-D commodity graphics hardware performs bump mapping in a texture pipeline that includes dot products instructions.

[0105] In one implementation the 3-D commodity graphics hardware component of video adapter/accelerator 844 is intended to function as a 3-D graphics accelerator having a specialized rendering subsystem which is designed to off-load graphics processing functions from host processing unit 804, thereby improving system performance. As such, an application program from any of program modules 828-830 executing on host processing unit 804 generates three-dimensional geometry data including information corresponding to points on the surface of a three-dimensional graphical object. These points are usable as vertices of polygons which, when connected, may be rendered to form a representation of the graphical object. The application program causes host processing unit 840 to transfer this 3-D geometry data to the graphics accelerator, along with corresponding control and connectivity information. The graphics accelerator receives this stream of 3-D geometry data and renders the encoded polygons on an attached display device.

[0106] Hosting processing unit 804 and video adapter/accelerator 844 function together as an example of a means for generating a bump mapped surface of a three-dimensional graphical object and a means for generating a horizon map of the bump mapped surface. System memory 806 is an example of a means for storing the generated horizon map in a texture map for reuse in rendering the three-dimensional graphical object. Host processor 804 and video adapter/accelerator 844 are illustrated as being coupled to monitor 842, several input devices, and system memory 806. This coupling enable control of bump and horizon map generation as well as the pixel representations in the frame buffer of system memory 806 to produce on the monitor 842 a rendering of shadows cast by light upon the bump mapped surface of the three-dimensional graphical object as a function of the horizon map.

[0107] In addition to the monitor 842, other output peripheral devices can include components such as speakers (not shown) and a printer 846 which can be connected to computer 802 via the input/output interfaces 840.

[0108] Computer 802 can operate in a networked environment using logical connections to one or more remote computers, such as a remote computing device 848. By way of example, the remote computing device 848 can be a personal computer, portable computer, a server, a router, a network computer, a peer device or other common network node, and the like.

[0109] The remote computing device 848 is illustrated as a portable computer that can include many or all of the elements and features described herein relative to computer 802. Logical connections between computer 802 and the remote computer 848 are depicted as a local area network (LAN) 850 and a general wide area network (WAN) 852. Such networking environments are commonplace in offices, enterprise-wide computer networks, intranets, and the Internet.

[0110] When implemented in a LAN networking environment, the computer 802 is connected to a local network 850 via a network interface or adapter 854. When implemented in a WAN networking environment, the computer 802 typically includes a modem 856 or other means for establishing communications over the wide network 852. The modem 856, which can be internal or external to computer 802, can be connected to the system bus 808 via the input/output interfaces 840 or other appropriate mechanisms. It is to be appreciated that the illustrated network connections are exemplary and that other means of establishing communication link(s) between the computers 802 and 848 can be employed.

[0111] In a networked environment, such as that illustrated with computing environment 800, program modules depicted relative to the computer 802, or portions thereof, may be

stored in a remote memory storage device. By way of example, remote application programs 858 reside on a memory device of remote computer 848. For purposes of illustration, application programs and other executable program components such as the operating system are illustrated herein as discrete blocks, although it is recognized that such programs and components reside at various times in different storage components of the computing device 802, and are executed by the data processor(s) of the computer.

[0112] Computer-Executable Instructions

An implementation of interactive horizon mapping may be described in the general context of computer-executable instructions, such as program modules, executed by one or more computers or other devices. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Typically, the functionality of the program modules may be combined or distributed as desired in various embodiments.

[0113] Exemplary Operating Environment

Figure 8 illustrates an example of a suitable operating environment 800 in which an exemplary interactive horizon mapping may be implemented. Specifically, the exemplary interactive horizon mapping described herein may be implemented (wholly or in part) by any program modules 828-830 and/or operating system 826 in Figure 8 or a portion thereof.

[0114] The operating environment is only an example of a suitable operating environment and is not intended to suggest any limitation as to the scope or use of functionality of the exemplary interactive horizon mapping described herein. Other well known computing systems, environments, and/or configurations that are suitable for use include, but are not limited to, personal computers (PCs), server computers, hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, programmable consumer

electronics, wireless phones and equipments, general- and special-purpose appliances, application-specific integrated circuits (ASICs), network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

[0115] Computer Readable Media

An implementation of an exemplary interactive horizon mapping may be stored on or transmitted across some form of computer readable media. Computer readable media can be any available media that can be accessed by a computer. By way of example, and not limitation, computer readable media may comprise “computer storage media” and “communications media.”

[0116] “Computer storage media” include volatile and non-volatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules, or other data.

Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by a computer.

[0117] “Communication media” typically embodies computer readable instructions, data structures, program modules, or other data in a modulated data signal, such as carrier wave or other transport mechanism. Communication media also includes any information delivery media.

[0118] The term “modulated data signal” means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By

way of example, and not limitation, communication media includes wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, RF, infrared, and other wireless media. Combinations of any of the above are also included within the scope of computer readable media.

[0119] For purposes of the explanation, specific numbers, materials and configurations are set forth above in order to provide a thorough understanding of the present invention.

However, it will be apparent to one skilled in the art that the present invention may be practiced without the specific exemplary details. In other instances, well-known features are omitted or simplified to clarify the description of the exemplary implementations of present invention, and thereby better explain the present invention. Furthermore, for ease of understanding, certain method operations are delineated as separate operations; however, these separately delineated operations should not be construed as necessarily order dependent in their performance.

[0120] The present invention may be embodied in other specific forms without departing from its spirit or essential characteristics. The described embodiments are to be considered in all respects only as illustrative and not restrictive. The scope of the invention is, therefore, indicated by the appended claims rather than by the foregoing description. All changes which come within the meaning and range of equivalency of the claims are to be embraced within their scope.